**Deep Learning（DL）项目周报**

2016.10.16

## 与方舟和闽峰的讨论

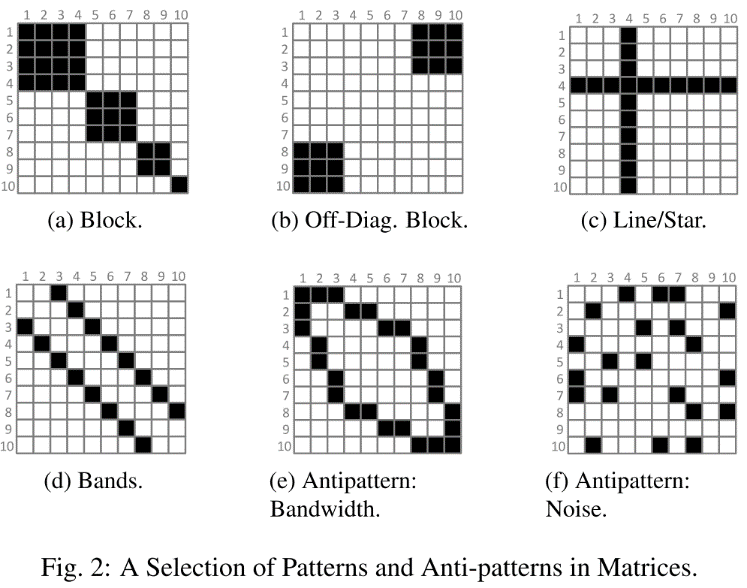
上周末与方舟和闽峰讨论了Graph Visualization方面怎样套用deep learning的问题，主要内容有以下方面：

* 关于graph layout
  + 无约束的graph layout方面，OpenOrd等方法已经可以实现百万级节点图小时级别完成布局，可以算作无约束布局算法里最快的方法。然而暂时没有想到好的方法，将基于弹簧模型/原子斥引力模型的布局算法建模为分类问题。
  + 在有约束布局方面，讨论了类似北大组几篇图可视化工作，以及现有其他工作（如正交直角graph）
  + 在应用方面，可以将手机基站数据当作一个应用，因为手机基站固定，即节点位置和数目固定。
* 关于更广阔的graph主题
  + 方舟推荐了一篇关于动态图中可视化任务的综述。接下来可能可以从任务角度出发，看哪些任务适合DL方法完成。

## 已有方向探索

#### Graph可视化

* VAST2016有一篇文章，提出了graph矩阵布局结果上的形状描述子Magnostics。这种描述子(descriptor)类似散点图的scagnostics，将矩阵布局中一些典型的几何形状特征进行分类，如下图所示。类似scagnostics，这种基于形状的识别匹配可能可以比较好地适配于DL方法。

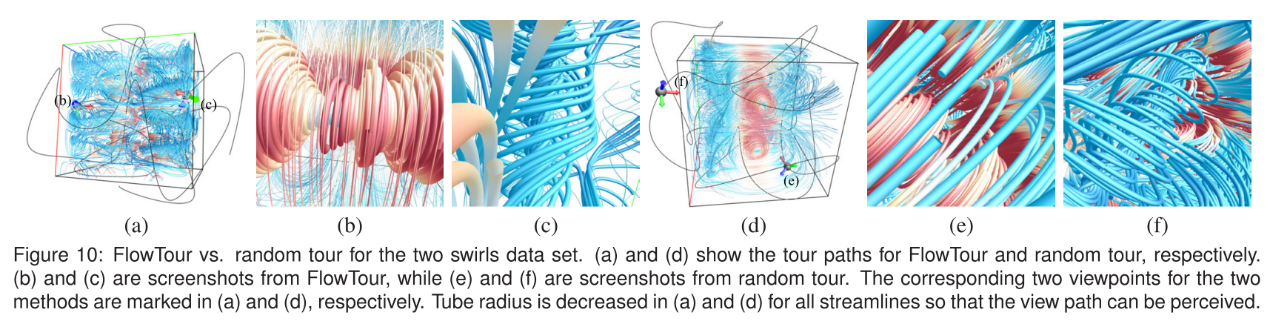


* 本周简单调研了数据挖掘方面关于graph classification的研究
  + 任务：根据有标记的graph结构，推断无标记的graph的分类标记。例如化学分子的毒性推断、社交网络的匹配等。
  + 方法
    - 定义一些graph的特征度量，然后将每个graph变换为以特征度量为维度的特征向量，使用特征向量进行图分类。
    - Kernel方法。主要是定义graph相似度的kernel，然后使用SVM等支持kernel的分类方法进行图分类。
* Deep learning方面，本周找到一篇slides，讲了目前如何使用DL方法解决graph mining的问题，主要包括：
  + DeepWalk：使用LSTM等时序神经网络方法模拟random walk。
  + 使用deep auto-encoder等方法，将graph转化为更好的特征表达。
  + 使用embedding方法，将graph结构嵌入至特征空间中，再用对应的特征向量放入DNN中。

#### 流场可视化

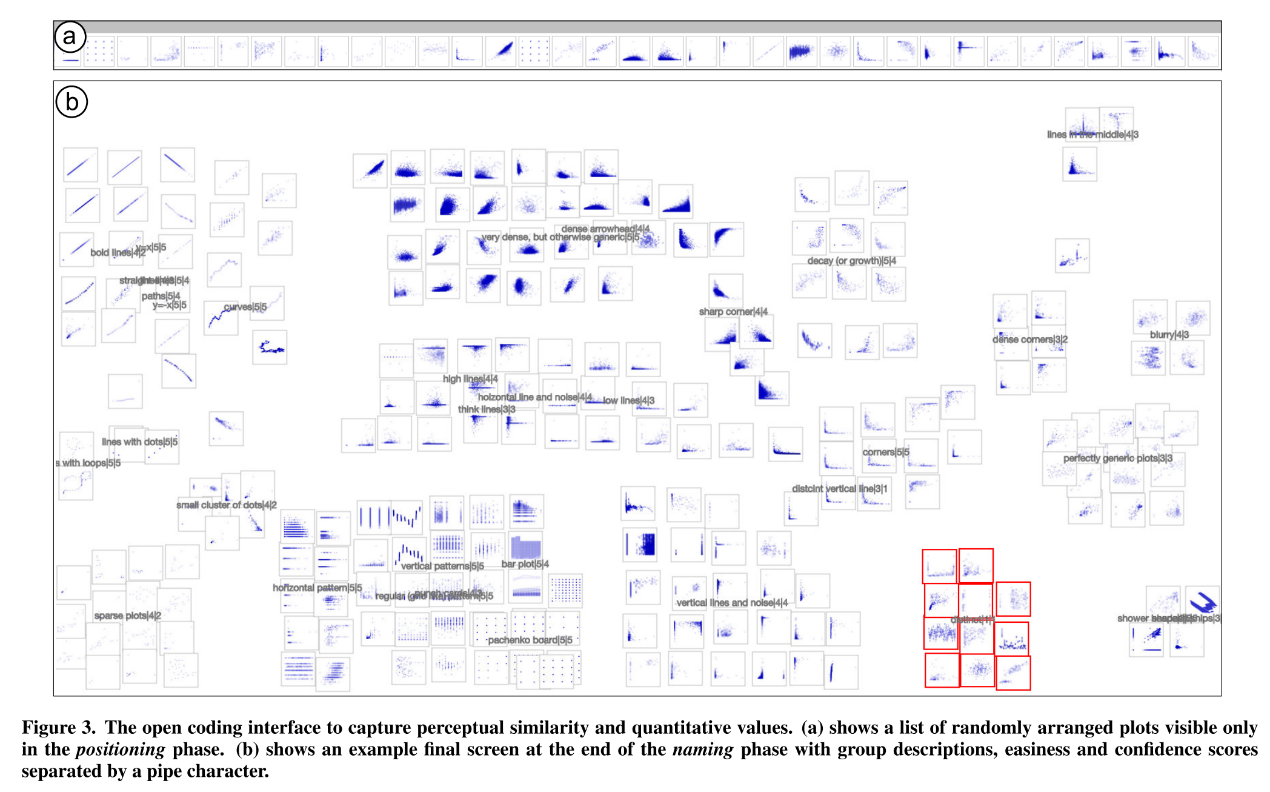
本周调研了IEEE近三年的流场可视化相关论文：

* 大部分文章的内容是提出新的视觉表达方法，或是解决一个之前的交互问题。
* PVis2014的文章FlowTour对DL项目有一定参考。该文章提出了一种视角选择方法，在三维流场可视化结果中自动计算出一条浏览路线及视角，以更好地表达流场中的特征。



#### 散点图

* 本周看到一篇CHI 2016上Enrico Bertini组的文章，调查了人类用户如何提取散点图上的特征，并对散点图进行分类。该文章构建了一组用户调研，让用户对247张散点图做分类，并让用户标注每个类别的特征和分类理由。文章发现现有的scagnostics等散点图特征描述子无法很好地捕捉到人类用户的分类标准和特征，并提出今后地描述子定义应更多从用户感知出发，而不是几何或数据特征定义。



某用户的自主分类结果，排列在一起的散点图被认作是同一类别的散点图。其中，每个类别的group上都用文字标签标明了该类别的特征。

## 一些想法与下一步计划

#### 想法

* 关于如何提出一种分析套路
  + Scagnostics、Magnostics等方法都是基于几何特征或某种标准制定了可视化结果中的特征描述子。我们可以从定义某种可视化形式的类别描述和特征描述子定义出发，针对该可视化形式提出一套（基于可视化结果图像的）特征描述方法和分类方法，进而使用DL方法做分类。
  + 例如：graph layout中，我们可以提出可视化结果上哪些特征是非常重要的，对应着某种分析任务，进而使用DL方法进行特征分类。
  + 我们的目标任务的可视化结果数据量应该非常庞大。例如实时动态图可视化或动态流场可视化，用户难以通过简单的动画等方法实时捕捉pattern，但DL方法可以当作用户的第二双眼帮助捕捉和筛选。

#### 下一步计划

接下来个人想法是从以下几条线继续跟进：

* 图可视化
  + Layout问题：之前跟方舟讨论的一个结论是，怎样将一个layout问题建模为分类问题。下周将以有约束的图布局问题为主，继续与方舟跟进这条线。
* （新增）轨迹可视化
  + 以基站数据为例，有以下思路：
    - 可以将基站之间的流量看作节点固定的动态图做分析。节点固定意味着比较适合作为DNN的输入。
    - 将可视化结果作为image进行特征识别。
  + 探索其他轨迹可视化形式是否也可使用DL方法做特征识别。
* 散点图
  + 先从现有特征描述子出发，尝试做一些实验。
* 流场：暂时先等领域专家的接入。
* 继续发散探索其他可视化方法中可能的任务。

## 引用

图可视化相关

Hu, Y. (2005). Efficient and High Quality Force-Directed Graph Drawing. The Mathematica Journal, 10(1), 37–71. http://doi.org/10.3888

Henaff, M., Bruna, J., & LeCun, Y. (2015). Deep Convolutional Networks on Graph-Structured Data. arXiv, 1–10. Retrieved from http://arxiv.org/abs/1506.05163

Behrisch, M., Bach, B., Hund, M., Delz, M., von Ruden, L., Fekete, J.-D., & Schreck, T. (2016). Magnostics: Image-based Search of Interesting Matrix Views for Guided Network Exploration. IEEE Transactions on Visualization and Computer Graphics, 1–1. http://doi.org/10.1109/TVCG.2016.2598467

流场可视化相关

Anwar, a., Nagel, T., & Ratti, C. (2014). Traffic Origins: A Simple Visualization Technique to Support Traffic Incident Analysis. 2014 IEEE Pacific Visualization Symposium, 316–319. http://doi.org/10.1109/PacificVis.2014.35

Battista, G. Di, Donato, V. Di, Patrignani, M., Pizzonia, M., Roselli, V., & Tamassia, R. (2015). Bitconeview: visualization of flows in the bitcoin transaction graph. In 2015 IEEE Symposium on Visualization for Cyber Security (VizSec) (pp. 1–8). IEEE. http://doi.org/10.1109/VIZSEC.2015.7312773

Chandler, J., Obermaier, H., & Joy, K. I. (2015). Interpolation-Based Pathline Tracing in Particle-Based Flow Visualization. IEEE Transactions on Visualization and Computer Graphics, 21(1), 68–80. http://doi.org/10.1109/TVCG.2014.2325043

Guo, D., & Zhu, X. (2014). Origin-destination flow data smoothing and mapping. IEEE Transactions on Visualization and Computer Graphics, 20(12), 2043–2052. http://doi.org/10.1109/TVCG.2014.2346271

Hong, F., Lai, C., Guo, H., Shen, E., Yuan, X., & Li, S. (2014). FLDA: Latent Dirichlet Allocation Based Unsteady Flow Analysis. IEEE Transactions on Visualization and Computer Graphics, 20(12), 2545–2554. http://doi.org/10.1109/TVCG.2014.2346416

Ibarra, J. C., Triana, J. A., & Hernández, J. T. (2016). Visualization of origin-destination matrices using a connection barchart and coordinated maps. IEEE Pacific Visualization Symposium, 2016–May, 189–193. http://doi.org/10.1109/PACIFICVIS.2016.7465268

Jun Ma, Walker, J., Chaoli Wang, Kuhl, S., & Ching Kuang Shene. (2014). FlowTour: An Automatic Guide for Exploring Internal Flow Features. In 2014 IEEE Pacific Visualization Symposium (pp. 25–32). IEEE. http://doi.org/10.1109/PacificVis.2014.14

Jun Tao, Chaoli Wang, Ching-Kuang Shene, & Seung Hyun Kim. (2014). A Deformation Framework for Focus+Context Flow Visualization. IEEE Transactions on Visualization and Computer Graphics, 20(1), 42–55. http://doi.org/10.1109/TVCG.2013.100

Kim, M., & Hansen, C. (2015). Surface flow visualization using the closest point embedding. 2015 IEEE Pacific Visualization Symposium (PacificVis), 17–23. http://doi.org/10.1109/PACIFICVIS.2015.7156351

Laramee, R. S., Hauser, H., Zhao, L., & Post, F. H. (2007). Topology-based flow visualization: the state of the art. Topology-Based Methods in Visualization, 1–19. http://doi.org/10.1007/978-3-540-70823-0

Matvienko, V., & Krüger, J. (2016). Explicit frequency control for high-quality texture-based flow visualization. 2015 IEEE Scientific Visualization Conference, SciVis 2015 - Proceedings, 41–48. http://doi.org/10.1109/SciVis.2015.7429490

Promrit, N., & Mingkhwan, A. (2015). Traffic Flow Classification and Visualization for Network Forensic Analysis. In 2015 IEEE 29th International Conference on Advanced Information Networking and Applications (pp. 358–364). IEEE. http://doi.org/10.1109/AINA.2015.207

Scheepens, R., Hurter, C., van de Wetering, H., & van Wijk, J. J. (2015). Visualization, Selection, and Analysis of Traffic Flows. IEEE Transactions on Visualization and Computer Graphics, 2626(c), 1–1. http://doi.org/10.1109/TVCG.2015.2467112

Tao, J., Wang, C. L., Shene, C. K., & Kim, S. H. (2014). A Deformation Framework for Focus plus Context Flow Visualization. Ieee Transactions on Visualization and Computer Graphics, 20(1), 42–55. http://doi.org/10.1109/tvcg.2013.100

Tao, J., Wang, C., Shene, C. K., Jun Tao, Chaoli Wang, & Ching Kuang Shene. (2014). FlowString: Partial Streamline Matching Using Shape Invariant Similarity Measure for Exploratory Flow Visualization. Visualization Symposium (PacificVis), 2014 IEEE Pacific, 9–16. http://doi.org/10.1109/PacificVis.2014.12

Tao, J., Wang, C., Shene, C.-K., & Shaw, R. A. (2016). A Vocabulary Approach to Partial Streamline Matching and Exploratory Flow Visualization. IEEE Transactions on Visualization and Computer Graphics, 22(5), 1503–1516. http://doi.org/10.1109/TVCG.2015.2440252

Yang, Y., Dwyer, T., Goodwin, S., & Marriott, K. (2016). Many-to-Many Geographically-Embedded Flow Visualisation: An Evaluation. IEEE Transactions on Visualization and Computer Graphics, 1–1. http://doi.org/10.1109/TVCG.2016.2598885

Zhang, J., Guo, H., & Yuan, X. (2016). Efficient unsteady flow visualization with high-order access dependencies. In 2016 IEEE Pacific Visualization Symposium (PacificVis) (pp. 80–87). IEEE. http://doi.org/10.1109/PACIFICVIS.2016.7465254

Zhang, J., Guo, H., & Yuan, X. (2015). High performance flow field visualization with high-order access dependencies. In 2015 IEEE Scientific Visualization Conference (SciVis) (pp. 165–166). IEEE. http://doi.org/10.1109/SciVis.2015.7429515

Tzeng, F. Y., & Ma, K. L. (2005). Intelligent feature extraction and tracking for visualizing large-scale 4P flow simulations. Proceedings of the ACM/IEEE 2005 Supercomputing Conference, SC’05, 2005(c). http://doi.org/10.1109/SC.2005.37

散点图相关

Pandey, A. V., Krause, J., Felix, C., Boy, J., & Bertini, E. (2016). Towards Understanding Human Similarity Perception in the Analysis of Large Sets of Scatter Plots. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI ’16, 3659–3669. http://doi.org/10.1145/2858036.2858155